31 ATGGCCAACAAGCTGAGGAAATCGAACGCCATCGAATGGGCCACCGGCCACCGGCACAGTACCGCTCCTGGAAAGGAGCTGCTGCCACAGC 1 M A N K L R K S N A I E W A T A T G T V P L L E R S C C H S 91 GAGGACGCCGCACTGGAGCCCCAAGCGAGCAAAACCAGCCATAGAGAACAAGCCCCCATCCTGCGCCACCTGAGCCAACTGAGCCACCTG 31 E D A A L E P Q A S K T S H R E Q A P I L R H L S Q L S H L 61 L I I A G L L I V C L A G V T E G R R H A P L M F E E S D T  ${\tt 271} \quad {\tt GGCAGGCGGTCCAACCGACCGGCGTCACCGAA\underline{TGCCAGTTTGGCAAAGTTTTGCGCGAA\underline{TTGGG}GTCCACCTGGTATGCGGATTTGGGT}$ 91 G R R S N R P A V T E C Q F G K V L R E L G S T W Y A D L G 361  $\tt CCACCCTTCGGAGTTATGTACTGCATCAAGTGTGAATGTGTGCGCATACCCAAGAAGCGGCGCATCGTTGCACGCGTCCAGTGTCGCAAT$ -CR1 PPFGVMYCIKCECVAIPKKRRIVARVQCRN ATCAAAAACGAGTGCCCGCCGGCCAAATGCGATGATCCCATCTCGTTGCCCGGAAAATGCTGCAAGACCTGTCCCGGCGATCGAAACGAT

I K N E C P P A K C D D P I S L P G K C C K T C P G D R N D 451 151 541 ACGGATGTAGCCTTGGATGTGCCCGTGCCCAATGAAGAGGAAGAGCGCAACATGAAACATTACGCTGCGTTGCTAACGGGCCGCACCTCC T D V A L D V P V P N E E E E R N M K H Y A A L L T G R T S 631 TATTTCCTCAAGGGTGAGGAAATGAAGTCCATGTACACCACCTACAATCCGCAGAATCTGGTGGCCACCGCCCGTTTCCTGTTCCACAAG 211 Y F L K G E E M K S M Y T T Y N P Q N V V A T A R F L F H K 721 AAGAATCTATACTACTCCTTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGTGTAATCCTGGAG 241 K N L Y Y S F Y T S S R I G R P R A I Q F V D D A G V I L E 811 GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCTGA 271 E H Q L E T T L A G T L S V Y A  $\bf N$  A  $\bf T$  G K I stop

FIGURE 1

1	ΑT	GGC	CAA	CAA	GCT	'GAG	GAA	ATC	GAA	CGC	CAT	'CGA	ATG	GGC	CAC	GGC	CAC	CGG	CAC	AGT	'ACC	GCT	'CCI	'GGA	AAG	GAG	CTG	CTO	CCA	CAGC	
1	M	A	N	K	L	R	K	S	N	Α	I	E	W	Α	Т	Α	T	G	Т	V	P	L	L	E	R	S	С	С	H	S	
91	GΑ	GGA	.CGC	CGC	ACT	'GGA	.GCC	CCA	AGC	'GAG	CAA	AAC	CAG	CCA	TAG	AGA	ACA	AGC	:CCC	CAT	CCI	'GCG	CCA	CCI	'GAG	CCA	ACT	'GAC	CCA	CCTG	TM
31	E	D	Α	A	L	E	P	Q	Α	S	K	Т	S	Н	R	Ε	Q	Α	P	1	L	R	, H	· L	s,	`Q	L	S	Н	_ <u>L</u>	
181	CI	CAT	CAT	CGC	CGG	ACT	GCT	GAT	CGT	CTG	CTT	GGC	:GGG	CGI	GAC	GGA	GGG	CCG	CCG	GCA	TGC	GCC	GCI	'CAT	GTT	'CGA	GGA	GTC	CGA	CACG	
61	<u>L</u>	I	I	Α	G	L	L	<u> </u>	V	С	L	A	G	<u>v</u>	Т	E	G	R	R	H	A	P	L	M	F	E	E	S	D	T	
271	GG	CAG	GCG	GTC	CAA	.CCG	ACC	AGC	GGT		_											GGG	GTC	CAC		_			TTT	GGGT	
91	G	R	R	S	N	R	Ρ	Α	V	Т	E	C	Q	F	G	K	V	L	R	Ε	L	G	S	Т	W	Ý	A	D	L	G	
361	CC	ACC	CTT	CGG	AGT	ТАТ	GTA	CTG	CAT	'CAA	.GTG	TGA	ATG	TGT	'GCG	CAT	ACC	CAA	GAA	.GCG	GCG	CAT	cgi	TGC	ACG	CGT	CCA	GTC	TCG	CAAT	CRI
121	P	P	F	G	V	М	Y	С	I	К	C	E	С	V	A	I	P	K	K	R	R	I	V	А	R	V	Q	С	R	N	<b>SCK</b> 1
451	AI	'CAA	AAA	CGA	GTG	CCC	GCC	'GGC	CAA	ATG	CGA	TGA	TCC	CAT	CTC	GTT	GCC	CGG	AAA	ATG	CTG	CAA	GAC	CTG	TCC	CGG	CGA	TCG	AAA	CGAT	
151	I	К	N	Ε	С	Р	P	A	K	С	D	D	Р	Ι	S	L	Р	G	K	C	С	.K	Т	C	P	G	D	R	N	D	İ
																													*	*	)
541																											GGG			CTCC	
181	T	D	V	А	Ь	D	V	Р	V	Р	N	Ε	E	Е	Ε	R	N	М	K	Н	Y	A	A	L	L	Т	G	R	Т	S	
631	TA	TTT	ССТ	CAA	GGG	TGA	GGA	ААТ	GAA	GTC	САТ	GTA	CAC	'CAC	CTA	CAA	TCC	GCA	GAA	тст	GGT	GGC	CAC	CGC	CCG	ттт	CCT	GTT	CCA	CAAG	
211	Y	F	L	K	G	E	E	M	K	s	M	Y	Т	T	Y	N	P	Q	N	V	V	А	Т	Α	R	F	L	F	Н	K	
721	AA	.GAA	TCT	ATA	CTA	CTC	CTT	CTA	CAC	CTC	ATC	GCG	AAT	'CGG	TCG	TCC	GCG	TGC	CAT	TCA	ATT	CGT	TGA	TGA	TGC	GGG	TGT	AAT	CCT	GGAG	
241	K	N	L	Y	Y	S	F	Y	T	S	S	R	I	G	R	P	R	Α	I	Q	F	V	D	D	Α	G	V	I	L	E	
811	GA	.GCA	TCA	ACT	GGA	GAC	CAC	CTT	GGC	GGG	CAC	TCT	CAG	TGT	CTA	TCA	GAA	TGC	CAC	GGG	CAA	GAT	ÇGG	CCG	CGG	CTC	GAG	GGT	ACC	TCTA	)
271	E	Н	Q	L	E	T	T	L	Α	G	Т	L	S	V	Y	Α	N	A	T	G	K	I	G	R	G	Ş	R	V	P	L	pUAS
																	*	*	*												Sequences
901	GA	GGA	TCT	TTC	TGA	AGG.	AAC	CTT	ACT	TCT	GTC	GTG	TGA	CAT	AAT	TGG.	ACA.	AAC	TAC	CTA	CAG	AGA	TTT	AAA	GCT	CTA	A				

FIGURE 2

1	ΑT	GGC	CAA	CAA	AGCI	GAC	GAA	ATC	GAA	CGC	CAT	'CGA	ATG	GGC	CAC	:GGC	CAC	CGG	CAC	AGT	'ACC	GCT	CCI	'GGA	AAG	GAG	CTG	CTG	CCA	CAGC	!
1	M	Α	N	K	L	R	K	S	N	Α	I	Ε	W	Α	Т	A	Т	G	T	V	P	L	L	E	R	s	С	С	Н	S	
91	GA	GGA	.CGC	CGC	ACT	GGA	GCC	CCA	AGC	GAG	CAA	AAC	CAG	CCA	TAG	AGA	ACA	AGC	.ccc	CAT	CCI	'GCG	CCA	CCT	GAG	CCA	ACT	GAG	CCA	CCTG	TM
31	Ε	D	A	A	L	Ε	P	Q	A	S	K	T	S	Н	R	Е	Q	Α	P	Ι	L	R	Н	L	.S	Q	L	S	Н	_ <u>L</u>	
181	CI	CAT	CAT	CGC	CGG	ACT									GAC	GGA	GGG	CCG	CCG	GCA	TGC	GCC	GCT	CAT	GTT	'CGA	GGA	GTC	CGA	CACG	
61	L	Ι	I	A	G	Ļ	L	I	V	С	L	A	G		Т	E	G	R	R	Н	Α	Ρ.	r P	M	F	. E.	E	S	D	Т	
271	GG	CAG	GCG	GTC	CAA	CCG	ACC	AGC	GGT	CAC	CGA	ATG	CCA	GTT	TGG	CAA	AGT	TTT	GCG	CGA	ATT	GGG	GTC	CAC	CTG	GTA	TGC	GGA	TTT	GGGT	)
91	G	R	R	S	N	R	P	A	V	T	E	С	Q	F	G	K	V	L	R	Е	L	G	Ş	Т	W	Y	Α.	D	L	G	
361	CC	ACC	CTT	'CGG	AGT	TAT	GTA	.CTG	CAT	CAA	.GTG	TGA	ATG	TGT	GCG	CAT.	ACC	CAA	GAA	GCG	GCG	CAT	CGT	TGC	NCG	CGT	CCA	GTG	TCG	CAAT	CR1:W->A
121	P	P	F	G	V	М	Y	Ç	Ι	K	C	E	С	V	Α	Ι	P	K	K	R	R	Ι	V	A	R	V	Q	С	R	N	SCRI:W->A
451	AT	CAA	AAA	CGA	GTG	CCC	GCC	GGC	CAA	ATG	CGA	TGA	TCC	CAT	CTC	GTT	GCC	CGG	AAA	ATG	CTG	CAA	GAC	CTG	TCC	CGG	CGA'	TCG	AAA	CGAT	
151	I	K	N	Ε	С	P	Р	A	K	С	D	D	P	1	S	L	Р	G	K	С	С	K	Т	С	P	G	D	R	N	D	]
																													*	* /	,
541 181									CGT	GCC	CAA			.GGA	AGA											_				CTCC	
191	*	D	٧	А	ы	D	V	Р	٧	Р	N	Е	Е	E	E	R	N	М	K	н	Y	A	A	L	L	Т	G	R	Т	S	
631	TA	TTT	CCT	CAA	.GGG	TGA	GGA	AAT	GAA	GTC	CAT	GTA	CAC	CAC	CTA	CAA'	TCC	GCA	GAA'	TCT	GGT	GGC	CAC	CGC	CCG	TTT	CCT	GTT	CCA	CAAG	
211	Y	F	L	K	G	E	E	M	K	S	M	Y	Т	Т	Y	N	P	Q	N	V	V	Α	T	A	R	F	L	F	Н	K	
721	AΑ	GAA	TCT	ATA	.CTA	CTC	CTT	CTA	CAC	CTC.	ATC	GCG	AAT	CGG'	TCG	TCC	GCG'	TGC	CAT'	TCA.	ATT	CGT	TGA	TGA'	TGC	GGG	TGT	<b>LA</b> A	ССТ	GGAG	
241	K	N	L	Y	Y	S	F	Y	Т	S	S	R	I	G	R	P	R	A	I	Q	F	V	D	D	A	G	V	I	L	E	
811	GA	GCA	TCA	ACT	GGA	GAC	CAC	CTT	GGC	GGG	CAC'	тст	CAG	TGT	CTA'	TCA	GAA'	TGC	CAC	GGG	CAA	GAT	ÇGG	CCG	CGG	CTC	GAG	GT	ACC	TCTA	)
271	Ε	Н	Q	L	E	T	Т	L	Α	G	Т	L	S	V	Y	Α	N	A	T	G	K	I	G	R	G	s	R	V	Р	L	<u>pUAS</u>
901	GΑ	GGA'	TCT	TTG	TGA	AGG.	AAC	CTT.	ACT	TCT	GTG	GTG	TGA	CAT	AAT'	TGG	* ACA	* AAC'	* TAC	CTA	CAG.	AGA'	TTT.	AAA(	GCT	CTA	A				Sequences
201	E.	n	т		다		T	Ŧ	T	T	Tall		n	Ŧ	т	C	NT.	т	T	3/	D	n	т	17	7	ا دسا	٦D				

FIGURE 3

1 M A N K L R K S N A I E W A T A T G T V P L L E R S C C H S  91 GAGGACGCCGCACTGGAGCCCCAAGCGAGCAAAACCAGCCATAGAGAACAAGCCCCCATCCTGGCCCACCTGGCCAACTGAGCCACCTG 31 E D A A L E P Q A S K T S H R E Q A P I L R H L S Q L S H L  181 CTCATCATCGCCGGACTGCTGATCGTCTGCTTGGCGGGCG	1	ΑT	GGC	CAA	.CAA	GCT	'GAG	GAA	ATC	'GAA	CGC	CAT	CGA	ATG	GGC	CAC	:GGC	CAC	CGG	CAC	AGT	ACC	GCT	CCT	GGA	AAG	GAG	CTC	CTG	CCA	CAG	2
18	1	M	A	N	K	L	R	K	S	N	A	I	E	W	A	T	A	Т	G	Т	V	₽	L	L	E	R	S	С	С	Н	S	
18	91	GA	GGA	CGC	CGC	ACT	GGA	AGCC	CCA	AGC	:GAG	CAA	AAC	CAG	CCA	TAC	AGA	ACA	AGC	ccc	CAT	ССТ	GCG	CCA	ССТ	'GAG	CCA	АСТ	GAG	CCA	ССТС	: TM
61 L I I A G L L I V C L A G V T E G R R H A P L M F E E S D T  271 GGCAGGGGGTCAACGGACCACGGGCGCCACGGATGCCAGTTTGGCAAAGTTTTGGCGCAATTGGGGTCCACCTGGTATGCGGATTTGGGT 91 G R R S N R P A V T E C Q F G K V L R E L G S T W Y A D L G  361 CCACCCTTCGGGGTTATGTACTGCCACTGGTGGAATGGTGGGCCACTGGTGGCGAATGGGGGCGCATCGTTGCACGGGTTGCACGGATTTGGGATT  451 P P F G V M Y C I K C B C V A I P K K R R I V A R V Q C R N  451 ACGAATGTAGCCTGGCGGCCAATGCCAGTGAATGCCAGTGTGGAATGGTGGGCCAATGGGGGGCAATGGTGCAGATGTGGGGCAATGGGATTGGGGCAATGGGATTGGGGCAATGGGATTGGGGCAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAAACGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAATGGAATGAAACAATAACGAATGGAATGAAT																																. 1111
61 L I I A G L L I V C L A G V T E G R R H A P L M F E E S D T  271 GGCAGGGGGTCAACGGACCACGGGCGCCACGGATGCCAGTTTGGCAAAGTTTTGGCGCAATTGGGGTCCACCTGGTATGCGGATTTGGGT 91 G R R S N R P A V T E C Q F G K V L R E L G S T W Y A D L G  361 CCACCCTTCGGGGTTATGTACTGCCACTGGTGGAATGGTGGGCCACTGGTGGCGAATGGGGGCGCATCGTTGCACGGGTTGCACGGATTTGGGATT  451 P P F G V M Y C I K C B C V A I P K K R R I V A R V Q C R N  451 ACGAATGTAGCCTGGCGGCCAATGCCAGTGAATGCCAGTGTGGAATGGTGGGCCAATGGGGGGCAATGGTGCAGATGTGGGGCAATGGGATTGGGGCAATGGGATTGGGGCAATGGGATTGGGGCAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAAACGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAATGGAATGAAACAATAACGAATGGAATGAAT	201	CIT.			000		<b>3</b> CC				ama	amm.										ma.a		~~~	~~~	-		~~-				
GCAGGCGGTCCAACCGACCAGCGGTCACCGAATGCCAATTGCGCAAAGTTTTGGGCAAATTGGGGTCCACCTGGTATGCGGATTTGGGT  91																_													_		CACC	<del>)</del>
91	-															-	_						-	_	••	-			-	-	-	`
CRI    CCACCCTTCGGAGTTATGTACTGCATCAAGTGTGAATGTGTGCGCATACCCAAGAAGCGGCGCATCGTTGCACGCGTGCCAGTGTCGCAAT   CCACCCTTCGGAGTTATGTACTGCATCAAGTGTGAATGTGTGCGCATACCCAAGAAGCGGCGCGCATCGTTGCACGCGTGCCAGTGTCGCAAT   CCACCCTTCGGAGTTATGTACTGCCGGCGCGCCAATGCGATGTGCAAGTGCGCGCGC											CAC	CGA	_															_				
ATCAAAAACGAGTGCCCGCCGGCCAAATGCGATGATCCCATCTCGTTGCCCGGAAAATGCTGCAAGACCTGTCCCGGCGATCGAAACGATCTCCTTCGTTGCCCGGAAAATGCTGCCAAGACCTGTCCCGGCGATCGAAACGATCTCCTTCCT	91	G	ĸ	R	S	N	R	Р	Α	٧	T	Е	C	Q	-F	G	K	V	L	R	E		G	S	T	W	- Y	- A		Ь	G	
ATCAAAAACGAGTGCCCGCCGGCCAAATGCGATGATCCCATCTCGTTGCCCGGAAAAATGCTGCAAGACCTGTCCCGGGATCGAAACGAT    ATCAAAAACGAGTGCCCCGCCGGCCAAATGCGATGATCCCATCTCGTTGCCCGGAAAATGCTGCAAGACCTGTCCCGGGATCGAAACGAT   151	361	CC	ACC	CTT	CGG.	AGT	TAT	GTA	.CTG	CAT	CAA	GTG	TGA	ATG	TGI	GCG	CAT	ACC	'CAA	.GAA	GCG	GCG	CAT	CGT	TGC	ACG	CGT	CCA	GTG	TCG	CAAT	CD1
151	121	Р	Р	F	G	V	М	Y	С	Ι	K	С	Е	С	V	A	I	P	K	K	R	R	Ι	₹ .	·A	R	V	Q	С	R	N	CKI
ACGGATGTAGCCTTGGATGTGCCCGTGCCCAATGAAGAGGGAAGAGCGCAACATGAAACATTACGCTGCGTTGCTAACGGGCCGCACCTCC  181	451	AT	'CAA	AAA	CGA	GTG	CCC	GCC	GGC	CAA	ATG	CGA	TGA	TCC	CAT	CTC	GTT	GCC	CGG	AAA	ATG	CTG	CAA	GAC	CTG	TCC	CGG	CGA	TCG	ΑΑΑ	ССАТ	,
181	151	Ι	K	N	E	C	P	P	Α	K	С	D	D	P	I	S	Ļ	P	G	K	C	С	K	Т	С	Р	G	D	R	N	D	
181	5 <i>4</i> 1	אכ	רכז	TOT	אממ	C THE	CCN	тст	ccc	ССТ	ccc	ሮ አ አ	<b>ጥር</b> እ	አ <i>ሮ</i> አ	CCA	አሮአ	ccc	רא א	<i>ር</i> አጥ	ר א א	ת יי	ת ידי א	ccc	TCC	ر بيس	CCT	7 7 C	ccc	ccc	* ~~~	* ·	,
211 Y F L K G E E M K S M Y T T Y N P Q N V V A T A R F L F H K  721 AAGAATCTATACTACTCCTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGGTGTAATCCTGGAG 241 K N L Y Y S F Y T S S R I G R P R A I Q F V D D A G V I L E  811 GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCGGCCGCGGCTCGAGGCAGGC																		_		_												
211 Y F L K G E E M K S M Y T T Y N P Q N V V A T A R F L F H K  721 AAGAATCTATACTACTCCTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGGTGTAATCCTGGAG 241 K N L Y Y S F Y T S S R I G R P R A I Q F V D D A G V I L E  811 GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCGGCCGCGGCTCGAGGCAGGC		*																														
AAGAATCTATACTACTCCTTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGTGTAATCCTGGAG 241 K N L Y Y S F Y T S S R I G R P R A I Q F V D D A G V I L E  811 GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCGGCCGCGGCTCGAGGCAGGC																					-											
241 K N L Y Y S F Y T S S R I G R P R A I Q F V D D A G V I L E  811 GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCGGCCGCGGCTCGAGGCAGCGCGGCC 271 E H Q L E T T L A G T L S V Y A N A T G K I G R G S R N R G  901 CGCATCTTTTACCCATACGATGTTCCTGACTATGCGGGCTATCCCTATCACGTCCCGGACTATGCACGATCCTATCATGACCTTCCA 301 R I F Y P Y D V F D Y A G Y P Y D V P D Y A G S Y P Y D V P  991 GATTACGCTGCTCAGTGCGGCCGCGATTATACGGACGACGACGACGACAAATCA	211	1	-	ם	IC	J	-		1-1	10	J	1.1	1	1	-	-	M	_	Q	14	٧	٧	Ω	1	^	K	Ľ	ם	r	п	K	
811 GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCGGCCGCGGCTCGAGGCAGCGCGGCCGCCGCCGCTCAGGCAGCACCACCGACCACCGACCACCGACCACCGACCACC			-	-																			CGT'	TGA'		TGC		TGT				
271 E H Q L E T T L A G T L S V Y A N A T G K I G R G S R N R G PUAS/PCR  901 CGCATCTTTTACCCATACGATGTTCCTGACTATGCGGGCTATCCCTATCACGTCCCGGACTATGCACGATCCTATCCATATGACCTTCCA 301 R I F Y P Y D V F D Y A G Y P Y D V P D Y A G S Y P Y D V P  991 GATTACGCTGCTCAGTGCGGCCGCGATTATACGGACGACGACGACAAATCA  [IX-FLAC	241	K	N	L	Y	Y	S	F	Y	Т	S	s	R	Ι	G	R	Р	R	A	Ι	Q	F	V	D	D	Α	G	V	Ι	L	Ε	
901 CGCATCTTTTACCCATACGATGTTCCTGACTATGCGGGCTATCCCTATCACGTCCCGGACTATGCACGATCCTATCCATATGACCTTCCA 301 R I F Y P Y D V F D Y A G Y P Y D V P D Y A G S Y P Y D V P  991 GATTACGCTGCTCAGTGCGGCCGCGATTATACGGACGACGACGACAAATCA	811	GA	GCA	TCA	ACT	GA.	GAC	CAC	СТТ	GGC	GGG	CAC	TCT	CAG	TGT	CTA	TCA	GAA	TGC	CAC	GGG	CAA	GAT	CGG	CCG	CGG	CTC	GAG	GCA	GCG	CGGC	LIA C/DCD
301 R I F Y P Y D V F D Y A G Y P Y D V P D Y A G S Y P Y D V P 991 GATTACGCTGCTCAGTGCGGCCGCGATTATACGGACGACGACGACAAATCA  SX-HA  SX-HA	271	E	Н	Q	L	E	T	T	L	Α	G	T	L	S	V	Y	Α	N	A	T	G	K	I	G	R	G	S	R	N	R	G	DUAS/PCK
301 R I F Y P Y D V F D Y A G Y P Y D V P D Y A G S Y P Y D V P 991 GATTACGCTGCTCAGTGCGGCCGCGATTATACGGACGACGACGACAAATCA  SX-HA  SX-HA	901	CG	САТ	СТТ	ттас	~CC.	ΔΤΔ	CGA	тст	ፐርር	тса	CTA	тсс	GGG	СТА	TCC	СΤΆ	* TC∆	* CGT	* ~~~	GGA	СТА	TGC	ACG:	בדר	CT2	TCC	מידם	тсь	درس	דרכם	`
991 GATTACGCTGCTCAGTGCGGCCGCGATTATACGGACGACGACGACAAATCA IX-FLAC													$\overline{}$							_			A									
	0.01	a.	mm *	999	Ta ar		ama	000	aaa	aa.				003	221		~~~															_
	331					N N	GTG C		CCG R	D	Y	K	D D				CAA.	ATC STC														

FIGURE 4

Sequence Range: 1 to 222

Nog protei	10 MDHSQCLVTIYAL	20 MVFLGLRIDQO	30 GGCQHYLHIRPA	40 APSENLPLVD	50 LIEHPDPIYDP	60 KEKDLNET	70 LLRTLMVGHFDP
SuperSog P [ 71 ] Nog protei					fvddagvIle-	-EhqLetT	280 LagTLsVyqn ^vv^^v^ ^^v LLRTLMVGHFDP
						A. 1. 34	•
	90 *	100	110	120	130	140	150
Nog protei	ILPEERLGVEDLG	ELDLLLRQKPS	GAMPAEIKGLE	EFYEGLQSKKI	HRLSKKLRRKL	QMWLWSQT	FCPVLYTWNDLG
SuperSog P [ 71 ]	300 IgrgsRvplEDLcl	EgtLLLw>	S Vector				
Nog protei	ILPEERLGVEDLG	ELDLLLR					
	170	180	190	200	210	220	
Nog protei	RYVKVGSCYSKRS	CSVPEGMVCKA	AKSMHLTILRW	VRCQRRVQQK	CAWITIQYPVI	SECKCSC	

FIGURE 5

ATATTAGATGGACATGCATAATAATTATTCATGTAACTATGTGATTTTCATTTTACACGAGGTGTAAGTCAGAATTTAAAATTCTTAAAA -261 TCATCAACAAGTATAAAATTGAAAATGTTAAAAATTGAAGCAGTCTTGGGTTTGCCTGCAACATGTTGCTGCCGATCGTTAGATQTTGCT 20.4 cDNA start -81 GCTGCATGTTGCCGCTGCATGTTGCCGTTGCATGTTGCCGCCGTTTGGCAACTTTATAAACACGGAGCGGATTCAGTCCTTCAGGTTCAG TCGCTCTTGAATTGCGTGGGATTGCACATCGGTCGTTCGGCTTTTTGGGTTCGGCATTTAGAGAGATACGATACGATCCGATACGATCCG 1.0 100 ATCCAGTACAAAAATCAAATTCAAACCGCACTCCCGATCCGGTCGCCATCATATACACGGCGGCTCGCACCGCAGCTATCTAGTAGATAA CGACTCTATTTTCCAGAGCAAGCTAAACACACTAGTGCTAAACCATAACTATATCTTAACTAAGGAAAACAAAGTCTCGAAACTGAAAA 280 370 460  $\tt TGTGTTATCTAATCTGCAAGAAGAAGTACAAGAATCGGGTATAGAATCGGCTCTATACTATATCTATACACCTGA TAPATCTATATCCATT$ 550 GTGTGTGCCAGTGTGTGCGTGCGACCTTTGTTTTTATATATTTTTTTGTTGTTGTTCATACTGTGAAACGTGCTTTTTACAAGCCGGTCG 730 1 AAGCTGAGGAAATCGAACGCCATCGAATGGGCCACGGCCACCGGCACAGTACCGCTCCTGGAAAGGAGCTGCTGCCACAGCGAGGACGCC 820 K L R K S N A I E W A T A T G T V P L L E R S C C H S E D A 910 GCACTGGAGCCCCAAGCGAGCAAAACCAGCCATAGAGAACAAGCCCCCATCCTGCGCCACCTGAGCCAACTGAGCCACCTGCTCATCATC TM A L E P Q A S K T S H R E Q A P I L R H L S Q L S H L L I 1000 GCCGGACTGCTGATCGTCTGCCTGGCGGGCGTGACGGAGGGCCGCCGCCATGCCCCCTCATGTTCGAGGAGTCCGACACGGGCAGGCGG L A G V T E G R R H A P L M F E E S D T G R R 1090 TCCAACCGACCAGCGGTCACCGAATGCCAGTTTGGCCAAAGTTTTGCGCGAATTGGGGTCCACCTGGTATGCGGATTTGGGTCCACCCTTC S N R P A V T E C Q F G K V L R E L G S T W Y A D L G P P F 1180 GGAGTTATGTACTGCATCAAGTGTGAATGTGTGGCGATACCCAAGAAGCGGCGCATCGTTGCACGCGTCCAGTGTCGCAATATCAAAAAC CR1 124 G V M Y C I K C E C V A I P K K R R I V A R V Q C R N I K N GAGTGCCCGCCGGCCAAATGCGATGATCCCATCTCGTTGCCCGGAAAATGCTGCAAGACCTGTCCCGGCGATCGAAACGATACGGATGTA ECPPAKCODPISLPGKCCKICPGDRNDTD GCCTTGGATGTCCCCGTGCCCAATGAAGAGGAAGAGGCGCAACATGAAACATTACGCTGCGTTGCTAACGGGCCCGCACCTCCTATTTCCTC 184 A L D V P V P N E E E E R N M K H Y A A L L T G R T S Y F L 1450 AAGGGTGAGGAAATGAAGTCCATGTACACCACCTACAATCCGCAGGAATGTGGTGGCCACCGCCCGTTTCCTGTTCCACAAGAAGAATCTA K G E E M K S M Y T T Y N P Q N V V A T A R F L F H K K N L TACTACTCCTTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGTGTAATCCTGGAGGAGCATCAA YYSFYTSSRIGRPRAIQFVDDAGVILEEHQ 1630 CTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCTGCGGTGTCTGGCGACGAGTTCCACGTGATTAC LETTLAGTLSVYQ**NAT**GKICGVWRRVPRDY K R I L R D D R L H V V L L W G N K Q Q A E L A L A G K V A 1810 K Y T A L Q T E L F S S L L E A P L P D G K T D P O L A G A GGTGGCACAGCGATCGTGTCCACCAGCAGCGGTGCCGCCTCATCGATGCATCTCACCCTGGTCTTCAATGGTGTCTTTGGTGCCGAGGAG 364 G G T A I V S T S S G A A S S M H L T L V F N G V F G A E E Y A D A A L S V K I E L A E R K E V I F D E I P R V R K P S 394 GCCGAGATCAATGTCCTGGAGCTGTCGTCGCCCATTTCCATACAGAATCTTCGACTGATGTCGCGTGGCAAACTCCTGCTGACCGTGGAG A E I N V L E L S S P I S I Q N L R L M S R G K L L L T V E 424 TCCAAGAAGTACCCACATCTGCGCATCCAGGGACACATCGTGACCCGAGCCAGCTGCGGAAATCTTCCAGACCCTGCTGGCGCACAGT 2170 S K K Y P H L R I Q G H I V T R A S C E I F Q T L L A P H S SR1 2260 GCCGAATCCTCGACCAAGAGCAGCGGTTTGGCGTGGGTCTACTTGAACACCGATGGATCTCTGGCCTACAACATCGAAACGGAGCACGTG A E S S T K S. S G L A W V Y L N T D G S L A Y N I E T E H V 2350 AACACCCGGGATAGGCCCAACATCAGTTTGATTGAGGAGCAGGGCAAGCGGAAGCTGGAGGATCTGACGCCGAGCTTCAACTTC NTRDRP**NIS**LIEEQG<u>KRK</u>AKLEDLTPSFNF 2440 AACCAGGCCATTGGTAGTGTGGAGAAGTTGGGTCCCAAGGTCCTCGAGTCGCTGTATGCCGGCGAACTGGGCGTTAATGTGGCCACCGAG

544 N Q A I G S V E K L G P K V L E S L Y A G E L G V N V A T E

2530 574	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	
2620 604	GAGCACACGGATGCACAGAATCCACATGCCGTCGGCATGGCCTGGATGTCCATTGACAACGAGTGCAATCTGCACTACGAGGTGACGCTC E H T D A Q N P H A V G M A W M S I D N E C N L H Y E ∜ T L	
2710 634	AACGGTGTGCCCGCCCAGGATCTGCAGCTGTATCTGGAGGAGAAGCCCATCGAGGCGATTGGAGCGCCAGTGACGAGGAAATTGCTCGAGNGVPAQDLQLYLEEKPIEAIGAPV, T.R.KLLESF	22
2800 664	GAATTCAACGGCTCCTATCTGGAAGGCTTCTTCCTCAGCATGCCATCCGCCGAACTGATCAAGCTGGAGATGAGCGTCTGCTATCTGGAG E F N G S Y L E G F F L S M P S A E L I K L E M S V C Y L E	
2890 694	GTCCATTCCAAGCACTCCAAACAGCTTCTGCTGCGCGGCAAACTGAAGAGCACCAAGGTGCCGGGTCACTGCTTCCCCGTCTATACGGAC V H S K H S K Q L L L R G K L K S T K V P G H C F P V Y T D	SR3
2980 724	AACAATGTTCCCGTGCCTGGCGACCACAATGATAACCATTTGGTGAACGGAGGACCAAGTGCTTTCACTCCGGACGCTTCTACAACGAA N N V P V P G D H N D N H L V N G E T K C F H S G R F Y N E	,,,,
3070 754	*	CR2
3160 784	CTCAAGTGCAAGTCCACGGAGCAACTGCTTCAGCGTGATGGTGAATGCTGTCCCAGCTGTTGTGCCCAAGAAGGAGGCCGCCGACTATTCA  L K C K S T E Q L L Q R D G E C C P S C V P K K E A A D Y S	
3250 814	GCGCAATCCTCGCCAGCCACCAATGCCACCGATTTGCTGCAACAGCGACGCGGCTGCCGCCTGGGCGAGCAGTTCCATCCCGCCGGTGCC  A Q S S P A T N A T D L L Q Q R R G C R L G E Q F H P A G A  * * *	
3340 844	S W H P F L P P N G F D T C T T C S C D P L T L E I R C P R	CR3
3430 874	CTCGTCTGCCCGCCGTTGCAGTGCAGCGAGAAGTGGGCCTATCGTCCAGACAAGAAGCATGCTGCAAGATCTGTCCGGAGGGCAAGCAG	
	L V C P P L Q C S E K L A Y R P D K K A C C K I C P E G K Q	
3520 904	AGCAGTTCCAATGGACACAAGACGACGCCGAACAATCCCAATGTGCTGCAGGATCAGGCCATGCAGCGATCGCCGAGTCACAGTGCCGAG S S S N G H K T T P N N P N V L Q D Q A M Q R S P S H S A E	
904 3610 934	AGCAGTTCCAATGGACACAAGACGACGCCGAACAATCCCAATGTGCTGCAGGATCAGGCCATGCAGCGATCGCCGAGTCACAGTGCCGAG S S S N G H K T T P N N P N V L Q D Q A M Q R S P S H S A E  GAGGTTCTGGCCAACGGCGGATGCAAGGTGGTCAACAAGGTGTACGAGAACGGCCAGGAGTGGCATCCGATCCTGATGTCCCACGGCGAG E V L A N G G C K V V N K V Y E N G Q E W H P I L M S H G E	
904 3610 934 3700 964	AGCAGTTCCAATGGACACAAGACGACGCCGAACAATCCCAATGTGCTGCAGGATCAGGCCATGCAGCGATCGCCGAGTCACAGTGCCGAG S S S N G H K T T P N N P N V L Q D Q A M Q R S P S H S A E  GAGGTTCTGGCCAACGGCGGATGCAAGGTGTCAACAAGGTGTACGAGAACGGCCAGGAGTGGCATCCGATCCTGATGTCCCACGGCGAG E V L A N G G C K V V N K V Y E N G Q E W H P I L M S H G E  CAGAAGTGCATCAAGTGCCGCTGCAAGGACTCCAAGGTGAACTGCGATGCCAAGCGCTGCTCCCGCTCCACGTGCCAGCAGCAGCACACGC Q K C I K C R C K D S K V N C D A K R C S R S T C Q Q Q T R	CR4
904 3610 934 3700 964 3790 994	AGCAGTTCCAATGGACACAAGACGACGCCGAACAATCCCAATGTGCTGCAGGATCAGGCCATGCAGCGATCGCCGAGTCACAGTGCCGAG S S S N G H K T T P N N P N V L Q D Q A M Q R S P S H S A E  GAGGTTCTGGCCAACGGCGGATGCAAGGTGTCAACAAGGTGTACGAGAACGGCCAGGAGTGGCATCCTGATGTCCCACGGCGAG E V L A N G G C K V V N K V Y E N G Q E W H P I L M S H G E  CAGAAGTGCATCAAGTGCCGCTGCAAGGACTCCAAGGTGAACTGCGATGCCAAGGGCTGCTCCCGCTCCACGTGCCAGCAGCAGCACACGC Q K C I K C R C K D S K V N C D A K R C S R S T C Q Q Q T R  GTGACCAGCAAACGGCGTCTGTTCGAGAAACCGGACGCAGCTGCTCCCGCCATCGATGAGTTCTGCTCCACCCAGTGCCGGAGATCGAGG V T S K R R L F E K P D A A A P A I D E F C S T Q C R R S R	ĽR4
3610 934 3700 964 3790 994 3880 1024	AGCAGTTCCAATGGACACAAGACGCCGAACAATCCCAATGTGCTGCAGGATCAGGCCATGCAGGCGATCGCCGAGTCACAGTGCCGAGS S S S N G H K T T P N N P N V L Q D Q A M Q R S P S H S A E  GAGGTTCTGGCCAACGGCGGATGCAAGGTGGTCAACAAGGTGTACGAGAACGGCCAGGAGTGGCATCCGATCCTGATGTCCCACGGCGAG E V L A N G G C K V V N K V Y E N G Q E W H P I L M S H G E  CAGAAGTGCATCAAGTGCCGCTGCAAGGACTCCAAGGTGAACTGCGATGCCAAGCGCTGCTCCCGCTCCACGTGCCAGCAGCAGCACACGC Q K C I K C R C K D S K V N C D A K R C S R S T C Q Q Q T R  GTGACCAGCAAAACGGCGTCTGTTCGAGAAACCGGACGCAGCTGCTCCGGCCATCGATGAGTTCTGCTCCACCCAGTGCCGAGATCGAGG CV T S K R R L F E K P D A A A P A I D E F C S T Q C R R S R  CGCCACCACAAGAGGGCAGCCGCATCATCAGCAGCGATCCTCCAGCTGACGGATGCCAGTGCCAGTTCAGTTCCA	ER4
904 3610 934 3700 964 3790 994 3880	AGCAGTTCCAATGGACACAAGACGACGCCGAACAATCCCAATGTGCTGCAGGATCAGGCCATGCAGCGATCGCCGAGTCACAGTGCCGAG S S S N G H K T T P N N P N V L Q D Q A M Q R S P S H S A E  GAGGTTCTGGCCAACGGCGGATGCAAGGTGTCAACAAGGTGTACGAGAACGGCCAGGAGTGGCATCCTGATGTCCCACGGCGAG E V L A N G G C K V V N K V Y E N G Q E W H P I L M S H G E  CAGAAGTGCATCAAGTGCCGCTGCAAGGACTCCAAGGTGAACTGCGATGCCAAGCGCTGCTCCCGCTCCACGTGCCAGCAGCAGCACACGC Q K C I K C R C K D S K V N C D A K R C S R S T C Q Q Q T R  GTGACCAGCAAACGGCGTCTGTTCGAGAAACCGGACGCAGCTGCTCCGGCCATCGATGAGTTCTGCTCCACCCAGTGCCGGAGATCGAGG V T S K R R L F E K P D A A A P A I D E F C S T Q C R R S R  CGCCACCACAAGAGGGCGCCGCATCATCAGCAGCGATCCTCCAGCTGACCGGTGCCCAGTTCCACCCATTCCATCAGTTCAGAT	CR4

FIGURE 6